

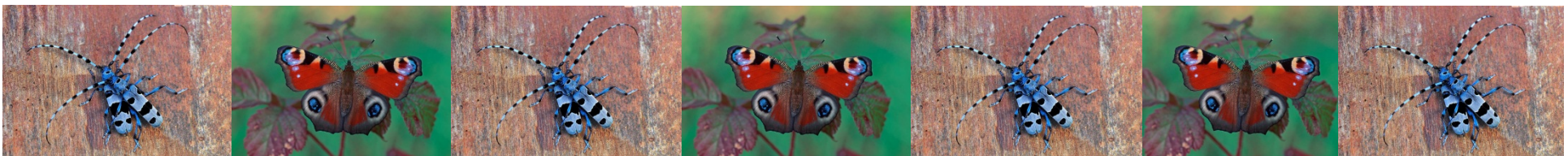


Introduction to N-mixture models

Short course

Marc Kéry

EURING Technical Meeting
Athens/GA, 28 April 2013



Preamble on occupancy and N-mixture models

- Occupancy model

Data (y_{ij})	Latent state (z_i)
0-1-0	1
1-0-1	1
1-1-1	1
0-0-0	0
0-0-0	1

State model: $z_i \sim \text{Bernoulli}(\psi_i)$

Observation model: $y_{ij} \sim \text{Bernoulli}(z_i p_{ij})$

- MacKenzie et al. (*Ecol.*, 2002); Tyre et al. (*Ecol. App.*, 2003)



Preamble on occupancy and N-mixture models

- “The” Nmix model

Data (y_{ij})	Latent state (N_i)
0-3-0	3
1-0-1	2
6-3-4	9
0-0-0	0
0-0-0	2

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

- Royle (*Biometrics*, 2004)



Outline of talk

- Introduction to Nmix family
- The “classical” Nmix model
- Assumptions and caveats
- Extensions and future directions



What are N-mixture models ?

- Class of hierarchical models with one submodel for measurement error (detection probability) and another for spatial and/or temporal variation in latent abundance states (N):

$$\begin{array}{ll} N_{it} \sim f(\lambda_i) & \# \text{ Model for abundance } N \\ C_{ijt} \sim g(N_{it}, \theta_{ijt}) & \# \text{ Measurement error model} \end{array}$$

- For variation of abundance (N), parametric (mixing) distribution is assumed -> N-mixture models
- Require data that are informative on measurement error for N, replicated in space (usually) and/or time (sometimes)
- Usually counts



N-mixture models as hierarchical models

- Hierarchical models (HM): nested sequence of random variables (observed or unobserved)

$$x \sim f(\omega)$$

$$y \sim g(x, \theta)$$

- e.g., randomised block ANCOVA (“mixed model”)

$$\alpha_i \sim \text{Normal}(\mu, \sigma_\alpha^2)$$

$$y_{ij} \sim \text{Normal}(\alpha_i + \beta * x_{ij}, \sigma^2)$$

- “The” N-mixture model (“explicit” HM)

$$N_i \sim \text{Poisson}(\lambda_i)$$

$$y_{ij} \sim \text{Binomial}(N_i, p_{ij})$$



Beauty and power of hierarchical models

- Can combine different pieces according to data collection protocol, modeling objectives, ...
- Nmix models for different observation protocols
- Family of N-mixture models
- by the way: terms “hierarchical model” and “state-space model” are synonymous to a large extent



Types of Nmix models: “the” Nmix model

- **Poisson/Binomial mixture model**

Data (y_{ij})	Latent state (N_i)	
-3-	3	
1-*-1	2	
6-3-*	9	
0-*-*	0	
-0-0	2	(denote NAs; pose no problems)

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

- Royle (*Biometrics*, 2004)



Types of Nmix models: Poisson/Bernoulli Nmix

- Data: detection/nondetection at site i during survey j

- Data (y_{ij}) Latent state (N_i)

0-1-0 3

1-0-1 2

1-1-1 9

0-0-0 0

0-0-0 2

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Bernoulli}(p_{ij})$

with $p_{ij} = 1 - (1-r_{ij})^{N_i}$

and r_{ij} = per-individual detectability



Royle & Nichols, *Ecology*, 2003 ("Royle-Nichols model")

vogelwarte.ch



Types of Nmix models: Poisson/Multinomial Nmix

- Many variants depending on data collection protocol, e.g., **removal sampling**
- Data: counts of “removals” in each time period j (= class k)

- Data (y_{ik}) Latent state (N)

5-1-0 8

3-1-2 9

1-1-0 5

0-0-0 0

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_i \sim \text{Multinom}(N_i, \pi_{ik})$

with $\pi_{ik} = f(p_{ij})$

- e.g., Royle, *Animal Biodiversity and Conservation*, 2004;
Dorazio et al., *Biometrics*, 2005



Types of Nmix models: Poisson/Multinomial Nmix

- **Capture-recapture or double-observer sampling**
- Data: # of each capture history k : e.g., 100, 010, 001, ...
- Data (y_{ik}) Latent state (N_i)

2-1-3-0-0-1-0	9
0-1-0-0-3-1-0	7
0-0-1-0-0-0-0	1

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_i \sim \text{Multinom}(N_i, \pi_{ik})$

with $\pi_{ik} = f(p_{ij})$

- e.g., Royle et al., *Ecol. Mono.*, 2007;
Webster et al, *JABES*, 2008



Types of Nmix models: Poisson/Multinomial Nmix

- Distance sampling (with binned distances)

- Data: Counts in each distance class k

- Data (y_{ik}) Latent state (N_i)

2-1-3 8

3-1-0 5

1-0-0 2

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_i \sim \text{Multinom}(N_i, \pi_{ik})$

with $\pi_{ik} = f(p_i)$

- e.g., Royle et al., *Ecol.*, 2004; Sillett et al, *Ecol.Appl.*, 2012



Types of Nmix models: Poisson/Poisson Nmix

- **Counts of animal cues etc.**
- Vector of counts, e.g., of fecal pellets, tracks along transect
- Data (y_{ij}) Latent state (N_i)

10-12-8	8
3-1-0	5
3-2-5	2

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Poisson}(N_i * \theta_{ij})$

- e.g., Stanley & Royle, *JWM*, 2005;
Guillera-Arroita et al., *JABES*, 2011



Other types of Nmix models

- Other published examples, by switching type of distribution for either abundance or detection
- see later for some alternative abundance models
- ditto for elaborations on detection model



Nmix models and occupancy models

- Seminal role of occupancy model of MacKenzie et al. (*Ecology*, 2002) and Tyre et al. (*Ecol.Appl.*, 2003)

- Historically:

site-occ (2002/3) -> Royle/Nichols (2003) -> Nmix (2004)

$$z_i \sim \text{Bern}(\psi_i)$$

$$N_i \sim \text{Pois}(\lambda_i)$$

$$N_i \sim \text{Pois}(\lambda_i)$$

$$y_{ij} \sim \text{Bern}(z_i * p_{ij})$$

$$y_{ij} \sim \text{Bern}(p_{ij})$$

$$y_{ij} \sim \text{Bin}(N_i, p_{ij})$$

- So, are Nmix models site-occupancy models ?



Are Nmix models site-occupancy models ?

- **Yes**, since any description of spatio-temporal patterns in abundance can be turned into a description in terms of occurrence/occupancy
- Occurrence (z) is deterministic function of abundance (N):
$$z = I(N > 0)$$
- Analogous with occupancy probability (ψ) :
$$\psi = \text{Prob}(N > 0)$$
- Occupancy: “the poor man’s abundance”
- see, e.g., Dorazio, *Ecology*, 2007



Are Nmix models site-occupancy models ?

- **No**, since Nmix models are not a special case of occupancy models
- Rather, both instances of “explicit” hierarchical models
- by “explicit” we mean that parameters have *explicit biological meaning*, e.g., abundance (N), occurrence (z)
- Unlike “expected abundance” or “expected occurrence” in many other hierarchical models for abundance or distribution
- (Calling Nmix models would be like calling all GLMs Probit regressions)



Should counts ever be degraded to det/nondet data ?

- Never !
- Only if absolutely have to, e.g. if assumptions of Nmix not warranted
- see later



An exercise in hierarchical modeling

- Re-invent the “classical” Nmix model from first principles
- most basic extension to model: adding covariates
- -> exercise on black board

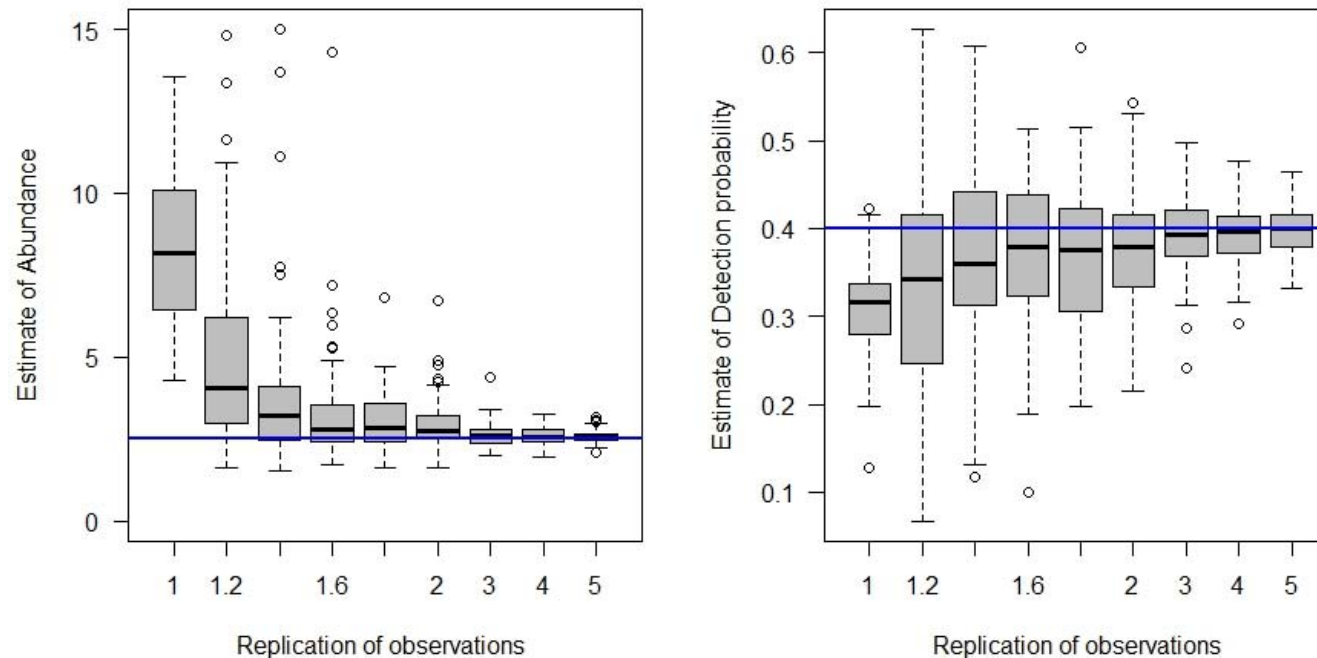


Different descriptions of Nmix model

- HM for counts replicated at R sites and T occasions with one mixture distribution (with param λ) for latent abundance states (N) and another distribution for observation process (with param p)
- “Explicit” HM: Parameters N have explicit biological meaning (*if* model well-specified)
- Nested GLM: Poisson GLM for N plus logistic regression as measurement error model
- Non-standard GLMM: logistic regression with nonstandard random effects (not normal, not continuous)



The need for replication



- but see work by Lele, Moreno, Solymos (fit Nmix to unreplicated data using penalized likelihood)
- Also F. Korner (unpublished ms)



Fitting of Nmix model

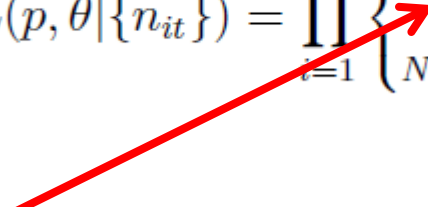
- Likelihood or Bayesian analysis
- Software: MARK, PRESENCE, R package **unmarked**, BUGS family, R, Matlab, PyMC,
- Likelihood analysis:

$$L(p, \theta | \{n_{it}\}) = \prod_{i=1}^R \left\{ \sum_{N_i = \max_t n_{it}}^{\infty} \left(\prod_{t=1}^T \text{Bin}(n_{it}; N_i, p) \right) f(N_i; \theta) \right\}$$



Fitting of Nmix model

- Likelihood or Bayesian analysis
- Software: MARK, PRESENCE, R package **unmarked**, BUGS family, R, Matlab, PyMC,
- Likelihood analysis:

$$L(p, \theta | \{n_{it}\}) = \prod_{i=1}^R \left\{ \sum_{N_i = \max_t n_{it}}^{\infty} \left(\prod_{t=1}^T \text{Bin}(n_{it}; N_i, p) \right) f(N_i; \theta) \right\}$$


- Infinite summation: replace infinity with reasonable upper summation limit (K) for N, e.g., 50, 100, ...
- Choose K such that likelihood for $N > K$ approx. 0



Likelihood analysis of Nmix model

- e.g., function `pcount()` in R package **unmarked**:

```
pcount(formula, data, K, mixture=c("P", "NB", "ZIP"),
starts, method="BFGS", se=TRUE, engine=c("C", "R"),
...)
```

- formula: R definition of linear models, e.g.,
~ 1 ~ 1 intercepts for p and N
~ wind ~ elev wind affects p, elevation N
- K: summation limit for likelihood evaluation: default
max(observed count) + 100
- mixture: Poisson, Negative binomial, Zero-inflated Poisson
(see later)



Likelihood analysis of Nmix model

- Advantages of likelihood analysis of Nmix model in **unmarked** (and MARK, PRESENCE):
 - get MLEs fast
 - linear model specification as usual in R
 - model selection using AIC or LRT
 - numerically reliable
- Disadvantages:
 - can't do nonstandard models (but most interesting data/models are nonstandard)
 - e.g., no random effects
 - no additional levels of hierarchy in HM
 - but see unmarked function `gpcount()`: fits TE emigration Nmix (Chandler et al., *Ecology*, 2011)



Bayesian analysis of Nmix model in BUGS family

- latent states N not removed by integration/summation
- N updated as part of MCMC sampling scheme
- specify hierarchical model almost exactly as written in algebra
- many extensions and nonstandard models trivial to code



Bayesian analysis of Nmix model in BUGS family

```
model {  
  # Priors  
  lambda ~ dunif(0, 50)  
  p ~ dunif(0, 1)  
  
  # Likelihood  
  for (i in 1:R) {  
    # True state model for the only partially observed true state  
    N[i] ~ dpois(lambda)      # True abundance state N at site i  
    for (j in 1:T) {  
      # Observation model for the actual observations  
      y[i,j] ~ dbin(p, N[i]) # Counts at i and j  
    }  
    z[i] <- step(N[i]-1)      # Occurrence indicator  
  }  
  # Derived quantities  
  total.N <- sum(N[])        # Total population size ar R sites  
  occ.fs <- mean(occ[])      # Finite sample occupancy  
}
```



Bayesian analysis of Nmix model

- Advantages:
 - model structure totally transparent (unlike in R)
 - usual advantages of Bayesian inference:
 - exact inference (no large-sample approximations)
 - random effects and other extensions trivial
 - estimates of latent variables (N) trivial; can do calculations on them
 - error propagation in derived quantities trivial (e.g., sum of N over R sites)
 - can introduce external information (informative priors)
- Disadvantages:
 - usual disadvantages of Bayesian inference: e.g., prior sensitivity
 - usual disadvantage of MCMC-based analysis: slow !
 - convergence assessment sometimes difficult



Benefits of Nmix model

- Conceptionally simple and plausible model
- Heart of model: Poisson GLM (we all know Poisson GLMs !)
- Estimate and model abundance (N) from “cheap” data
- “Cheap”: counts of unmarked individuals without individual identification (*)
- More data can be collected: e.g., more sites, more times, more temporal reps
- More information, e.g., about environmental relationships of abundance

(*) BUT see next slide !



Assumptions of Nmix model

- Closure: N_i constant over all surveys
- (Note closure assumption more severe than in site-occ)
- No individual ID: across occasions ID ignored
- But ID not ignored within occasion ! -> must exclude false-positives (double counts)
- N_i individuals detected independently
- All N_i individuals at occasion j have same detection probability p_{ij} (can only model p_{ij}): for instance, ignores effect of distance
- Parametric assumptions of model:
 - Poisson (with covariates, random effects etc.)
 - Binomial (with covariates, random effects etc.)



Test of assumptions

- Closure: this is a judgement:
 - Is study duration short relative to dynamics of system ?
 - Scale of movement of individuals relative to scale of sample plots (Efford & Dawson, *Ecosphere*, 2012)
- No-false-positives: similar judgment considerations, e.g., don't use model for (large) flocks
- Independent detections: ditto
- Homogeneity of detection (p_{ij}) and parametric assumptions:
 - Parametric bootstrap (likelihood analysis) GOF
 - Bayesian p-value GOF / posterior predictive checks (MCMC analysis)
 - in latter can test abundance and detection models separately (see p. 196 in Link & Barker, 2010)



Effects of assumptions violations

- Lack of closure: conventional wisdom: N_i refers to some superpopulation associated with sample plot
- can view as p-corrected index of per plot-abundance
- may be meaningless sometimes (i.e., when too much “temporary emigration”)
- not sure about false positives, independent detections, parametric model assumptions ?
- (but see Martin et al., MEE, 2011)



Remedies to assumption violations

- Closure:
 - design stage: make total study period short relative to system dynamics
 - analysis stage: discard some of the data; open models (see later); do occupancy modeling instead
- No individual ID: not so much
- Independence of detection: model non-independence (e.g., Martin et al., MEE, 2011; Dorazio et al., MEE, 2012)
- Homogeneity of p_{ij} : not much to be done
- Parametric assumptions of model (Poisson, Binomial, etc):
add complexity to model, e.g., covariates, random effects;
see later



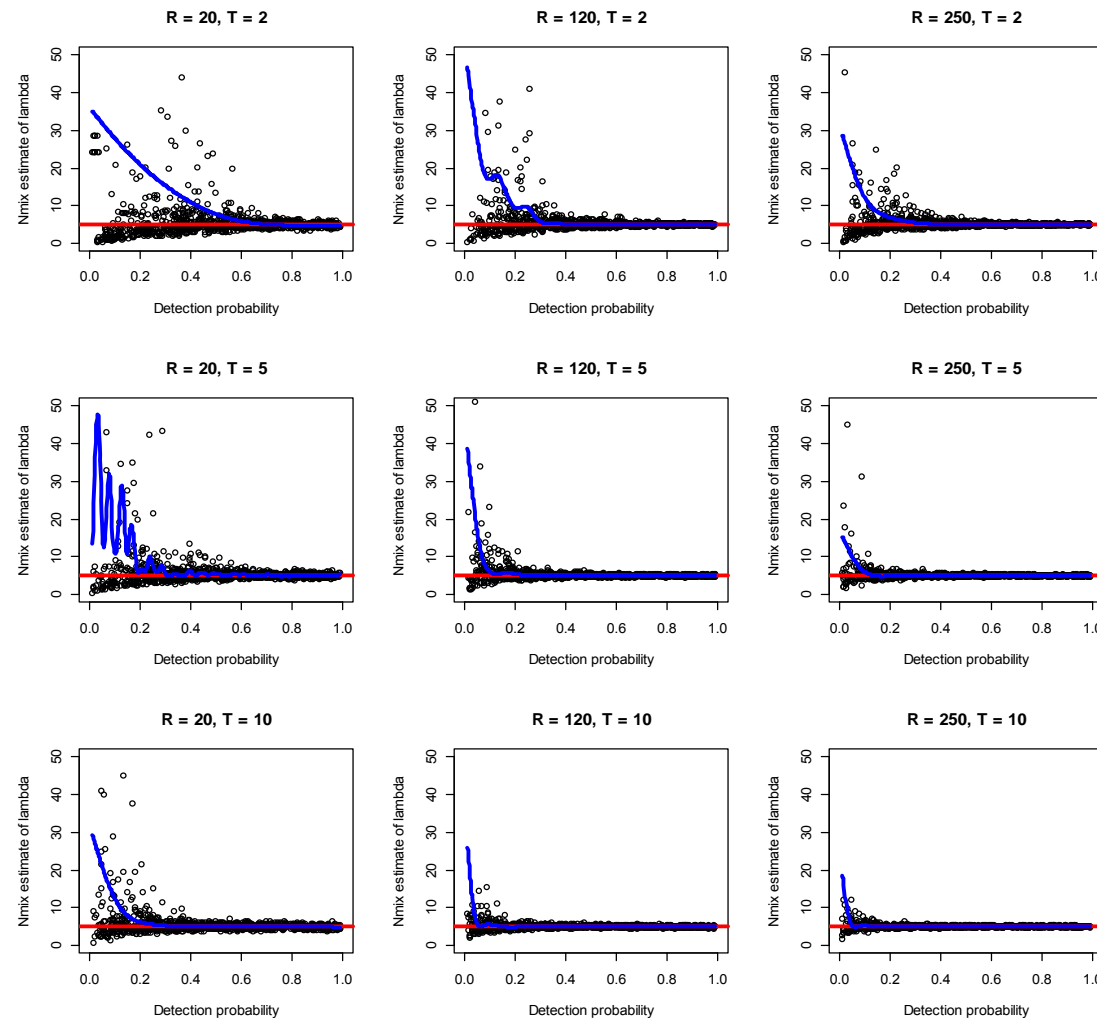
Identifiability problems in “the” Nmix ?

- Bill Link (unpublished ms):
 - intercept estimates highly correlated
 - with correlation 1, Nmix reduces to limiting case of Poisson model with random site effects
- Emily Dennis et al. (unpublished note): similar observations
- Couturier et al. (*JWM*, 2013): MLEs sensitive to choice of K , especially for small p
- problems particularly with small p
- hence, stay tuned for new findings, be wary with small p
- Possible remedies: Jack up p , use weakly informative priors (or constraints on K), collect extra data, use other member of Nmix family if can (or else Poisson random-effects model; see work by Link and Sauer on BBS analyses)



Use simulation to check quality of inferences

- trivial with program R, by varying #sites, #nreps, average N and p
- ex. MLEs



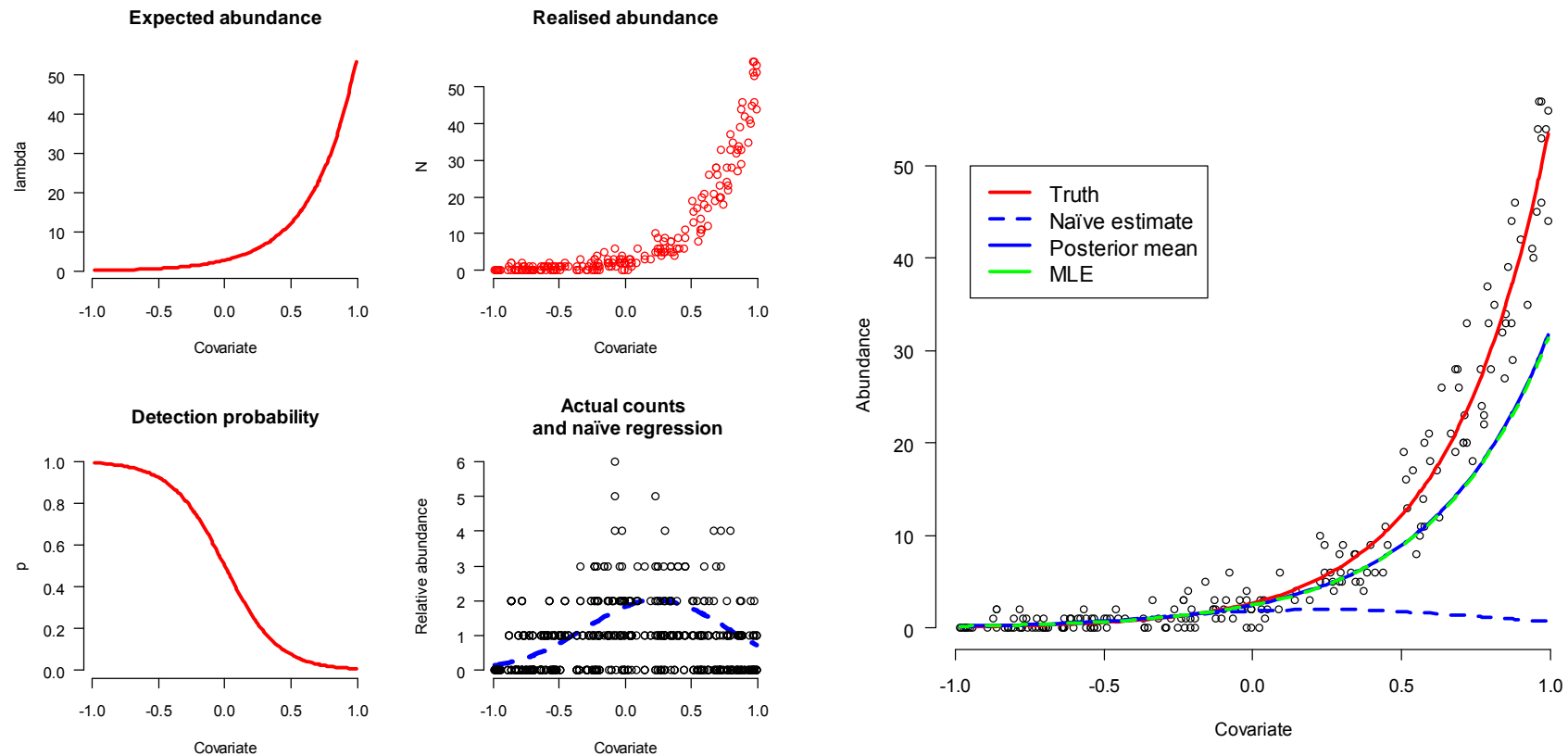
Two illustrations



vogelwarte.ch



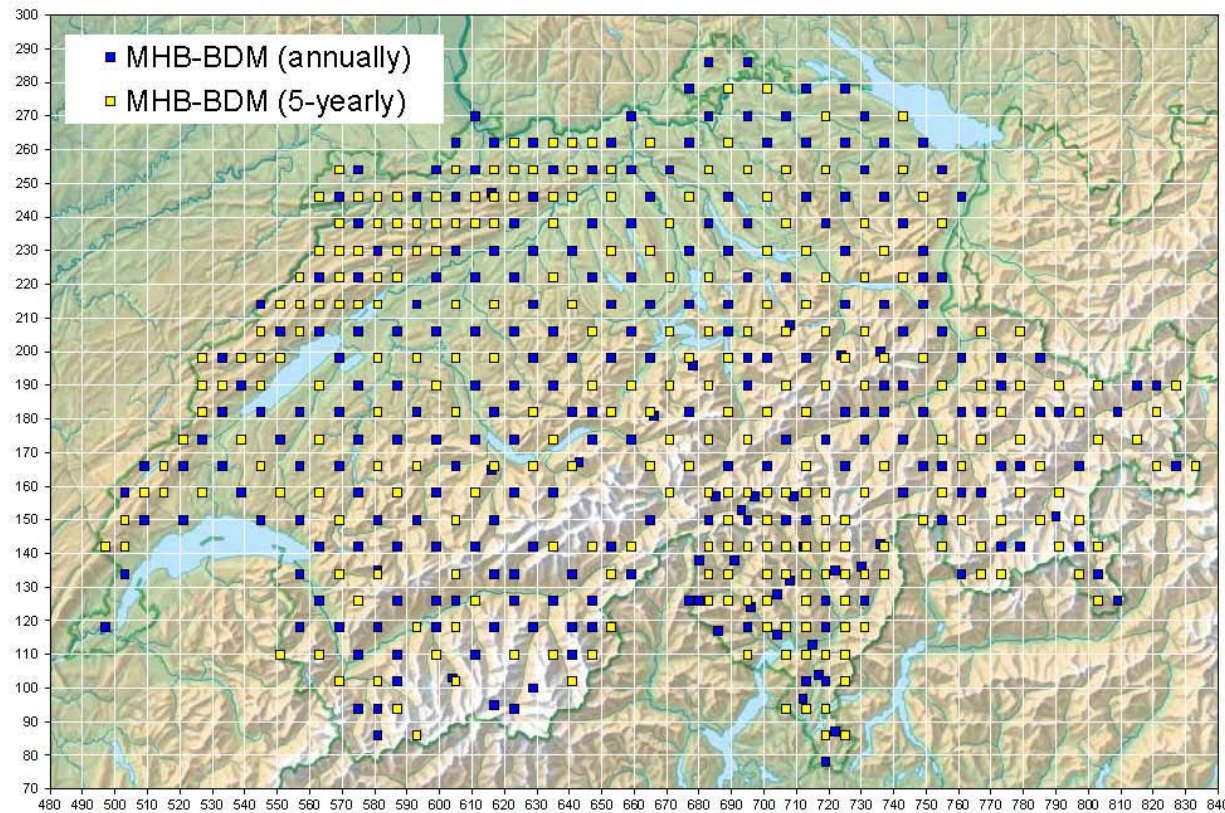
(A) Really tricky simulated data



- Importance of accounting for p !



(B) MHB: Swiss Breeding Bird survey



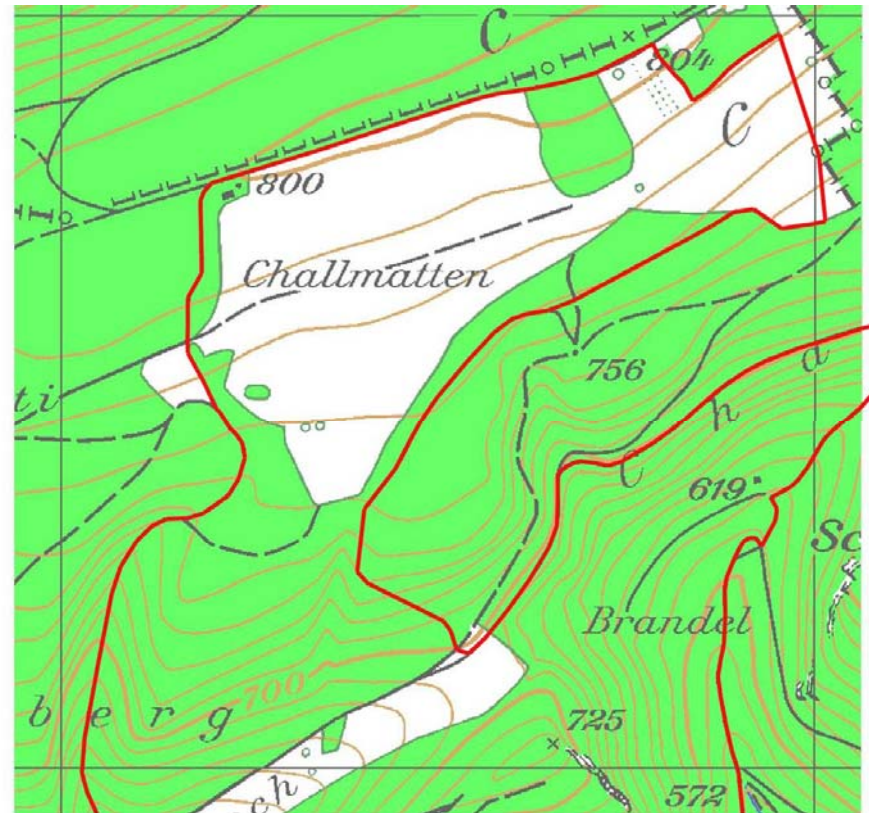
- 2 * 267 1km² quads, 3 reps/breeding season (15 April-30 June)



vogelwarte.ch



(B) MHB: Swiss Breeding Bird survey



- Territory mapping: record all locations of all species
- Here: reduce to counts for each survey



(B) MHB: Swiss Breeding Bird survey



- Andy Royle's favourite Swiss bird: the willow tit
- estimate and model abundance and map things
- MLEs from **unmarked**, model selection using AIC, parametric bootstrap GOF



(B) MHB: Swiss Breeding Bird survey

- e.g.,

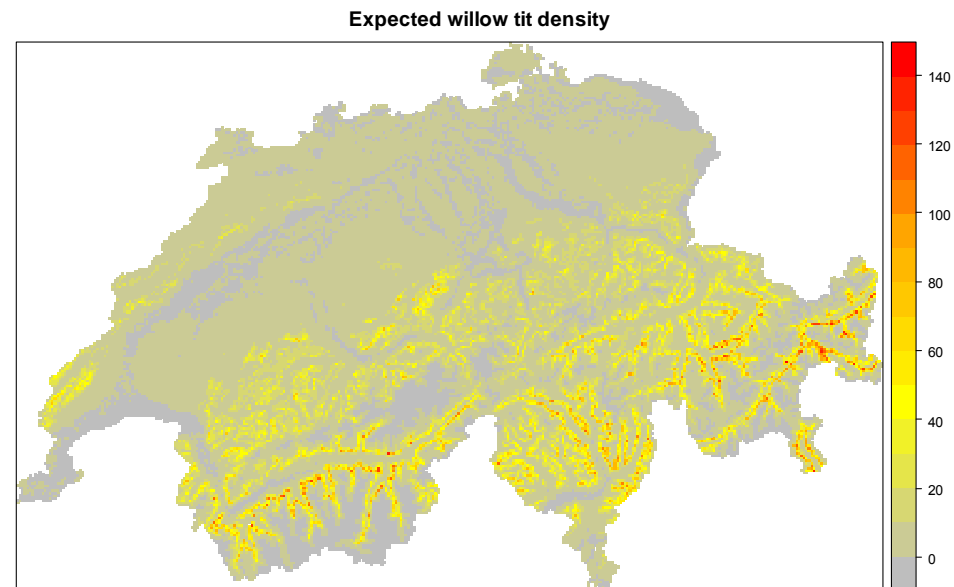
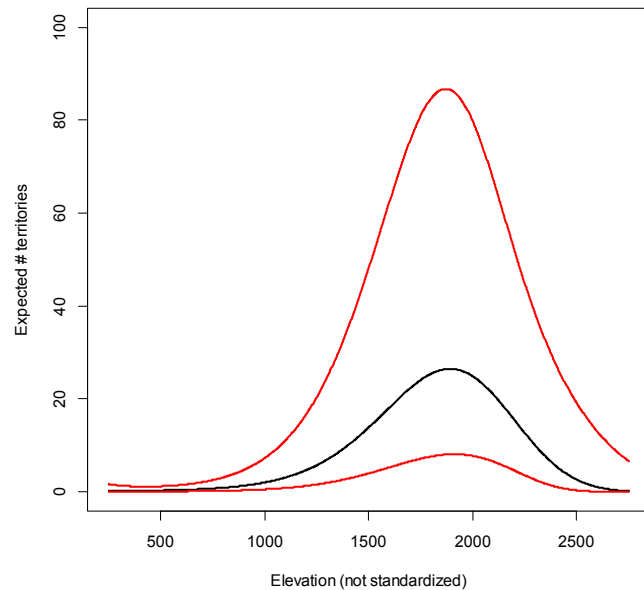
```
system.time(fm13 <- pcount(~day + (day^2)  
~forest+elev+I(elev^2)+ I(elev^3)+ length, mhb.umf))
```

- Best Poisson mixture model did not fit
- Negative Binomial (\approx overdispersed Poisson) did fit



(B) MHB: Swiss Breeding Bird survey

- what can we do after fitting the model ?
- -> try to understand what model is telling us: predictions !



- National population size estimate: $\sim 380,000$
- Bootstrap variance estimate: (132,110 - 1,193,516)



Important assumption ???



Space in Nmix and related models

- adding up plot-specific N estimates assumes plot area is known
- this is not usually strictly true: holes, edge territories
- would have to use Nmix with distance sampling or some sort of spatially explicit capture-recapture to relax the “known-area assumption”
- density and N can only be treated interchangeably when area known
- another assumption in many applications of Nmix models (or any kind of population size estimation)
- Chandler et al., *AOAS*, 2013: Nmix with underlying spatial model of animals



Time-for-space substitution in the Nmix

- instead of R spatial replicates in single season could have R temporal replicates (= seasons) of a single (or few sites)

- e.g., year observed data

1990	9-10-11
------	---------

1991	4-2-6
------	-------

...	
-----	--

2012	7-4-4
------	-------

- Seems to work well for >20 years; do simulations !
- example Yamaura et al., *JAPPL*, 2011 (community Nmix model)



Extensions of the model (for all Nmix family)

- add effects of measured covariates
- changes in abundance model (mixing distribution)
- changes in detection model
- add effects of unmeasured covariates (random effects)
- add space (e.g., spatial exponential correlation function, CAR random site effects)
- multiple species, abundance-based community models
- change closed model to open models:
 - “trend models”, Royle & Dorazio (2008);
 - implicit dynamics model, Chandler et al., *Ecology*, 2011
 - explicit dynamics model, Dail & Madsen, *Biometrics*, 2011



Extensions 1: covariates

- **Never forget:** heart of model is Poisson GLM, with logistic regression measurement error model attached
- All you can do with a Poisson or a Binomial GLM can also do to a Nmix model
- Adding covariates

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

- add log and logit linear regression models:

$$\log(\lambda_i) = \alpha_0 + \alpha_1 * x_{i,1}$$

$$\text{logit}(p_{ij}) = \beta_0 + \beta_1 * x_{i,1} + \beta_2 * x_{i,2} + \beta_3 * x_{ij,3}$$



Extensions 2: other abundance models

- Account for overdispersion: e.g., zero-inflation (1), Negative Binomial (2), Poisson log-normal (3):

1. State models: $z_i \sim \text{Bernoulli}(\psi_i)$

$$N_i \sim \text{Poisson}(z_i * \lambda_i)$$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

2. State model: $N_i \sim \text{Negative Binomial}(\lambda_i, \alpha)$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

3. State model: $N_i \sim \text{Poisson}(\lambda_i)$

$$\log(\lambda_i) = \alpha_i, \text{ with } \alpha_i \sim \text{Normal}(\mu_\alpha, \sigma)$$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$



Extensions 3: other detection models (overdisp.)

- Can account for effects of unobserved, latent covariates at site level (1), occasion level (2) or site-by-occasion level (3):

State model: $N_i \sim \text{Poisson}(\lambda_i)$

Observation model: $y_{ij} \sim \text{Binomial}(N_i, p_{ij})$

1. $\text{logit}(p_{ij}) = \beta_i$, with $\beta_i \sim \text{Normal}(\mu_\beta, \sigma)$
2. $\text{logit}(p_{ij}) = \beta_j$, with $\beta_j \sim \text{Normal}(\mu_\beta, \sigma)$
3. $\text{logit}(p_{ij}) = \beta_{ij}$, with $\beta_{ij} \sim \text{Normal}(\mu_\beta, \sigma)$



Extensions 4: adding space

- Standard models assume observations independent, given covariates
- Spatial or other dependencies may remain
- Add correlated, site-specific random effects :

$$N_i \sim \text{Poisson}(\lambda_i)$$

$$\log(\lambda_i) = \alpha_i, \text{ with } \alpha_i \sim \text{Normal}(\mu_\alpha, \Sigma)$$

Σ = distance-dependent variance-covariance matrix

- e.g., spatial-exponential correlation: Royle et al., *Ecol. Mono.*, 2007; Webster et al., *JABES*, 2008; Post van den Burg, *JAPPL*, 2011; Chelgren et al., *Ecology*, 2011
- conditional-autoregressive (CAR) models
- can both be implemented in WinBUGS/OpenBUGS (or MCMC coded up by hand)



Computationally expensive !



Extensions 5: multi-species (community) models

- joint Nmix model for all species observed in a community
- can use data-augmentation to estimate # species never seen (Royle et al., *JCGS*, 2007; Royle and Dorazio, 2008)
- express occurrence, and therefore species richness, as deterministic function of abundance
- usual advantage of random effects modeling: improved inferences for rare species (see e.g., Zipkin et al., *JAPPL*, 2009)
- examples: Yamaura et al., *JAPPL*, 2012; Chandler et al., *Conservation Biology*, in press



Extensions 6: open models

- three motivations:
 - account for closure assumption violation
 - estimate trends
 - explicitly estimate dynamics
- three models:
 - Dodd & Dorazio, *Herpetologica*, 2004;
Royle & Dorazio (book 2008);
Kéry et al., *JAPPL*, 2009;
Kéry & Schaub (book, 2012)
(treat years as a block, trend models)
 - Chandler et al., *Ecology*, 2011
(implicit dynamics model)
 - Dail & Madsen, *Biometrics*, 2011
(explicit dynamics model)



Treat years as a block approach (and trends)

- Fit separate parameters in abundance model for each year

$$N_{ik} \sim \text{Poisson}(\lambda_{ik}) \quad (k \text{ indexes years})$$

$$\log(\lambda_i) = \alpha_k + \text{stuff}$$

$$y_{ijk} \sim \text{Binomial}(N_i, p_{ij})$$

- can constrain annual estimates of $\log(\text{expected } N)$

$$\text{e.g., } \alpha_k = \alpha_0 + \beta * \text{year}_k$$

- β is trend parameter (see Royle & Dorazio, 2008;
Kéry et al., *JAPPL*, 2009)



Implicit dynamics: Chandler et al. (2011)

- multi-scale (3-level) model with one level for availability (1-temporary emigration)

Superpopulation model: $M_i \sim \text{Poisson}(\lambda)$
Random temporary emigration: $N_{ij} \sim \text{Binomial}(M_i, \theta)$
Observation model: $y_{ijk} \sim \text{Binomial}(N_i, p)$

- θ = Prob. of being exposed to sampling (1 – TE prob.)
- Assumes random temporary emigration described by θ
- implicit dynamics: random “in/out”
- fitting function in **unmarked**: `gpcount ()`



Explicit dynamics: the Dail-Madsen (2011) model

- explicit demographic model (population dynamics model)

Initial condition:	$N_{i1} \sim \text{Poisson}(\lambda)$
Survival process:	$S_{it} \sim \text{Binomial}(N_{it-1}, \omega)$
Recruitment process:	$G_{it} \sim \text{Poisson}(N_{it-1} * \gamma)$
Annual population size:	$N_{it} = S_{it} + G_{it}$
Observation model:	$y_{itk} \sim \text{Binomial}(N_{it}, p)$

- S_{it} : latent variable, survivors
- G_{it} : latent variable, recruits
- ω : apparent survival rate
- γ : recruitment rate



Explicit dynamics: the Dail-Madsen (2011) model

- Application: Chandler et al., *JAPPL*, 2011
 - mythical model: can estimate population dynamics from unmarked individuals
 - but: makes strong parametric assumptions
 - has produced unrealistic survival estimates
 - fitting function `pcountOpen()` in **unmarked**: VERY SLOW !
 - can be fit in JAGS, but not Win/OpenBUGS (no clue why)
 - more research is needed
-
- **2-3-year postdoc partly on this model available at Swiss Ornithological Institute RIGHT NOW**



Acknowledgements

- Andy Royle
- Richard Chandler
- Bill Link
- Emily Dennis
- Evan Cooch

